Combinations of SM Higgs searches at the LHC and Tevatron:

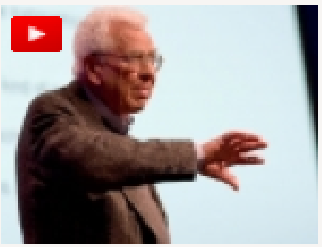
Information lost …
but not forgotten

Ben Kilminster
Fermilab

Zurich 2012 Higgs confrontation workshop
January 9, 2012

TED Ideas worth spreading

- Organization with > 1000 Short online talks on ideas worth spreading



Murray Gell-Mann on beauty and truth in physics
16:02 Posted: Dec 2007
Views **399,109** | Comments **112**

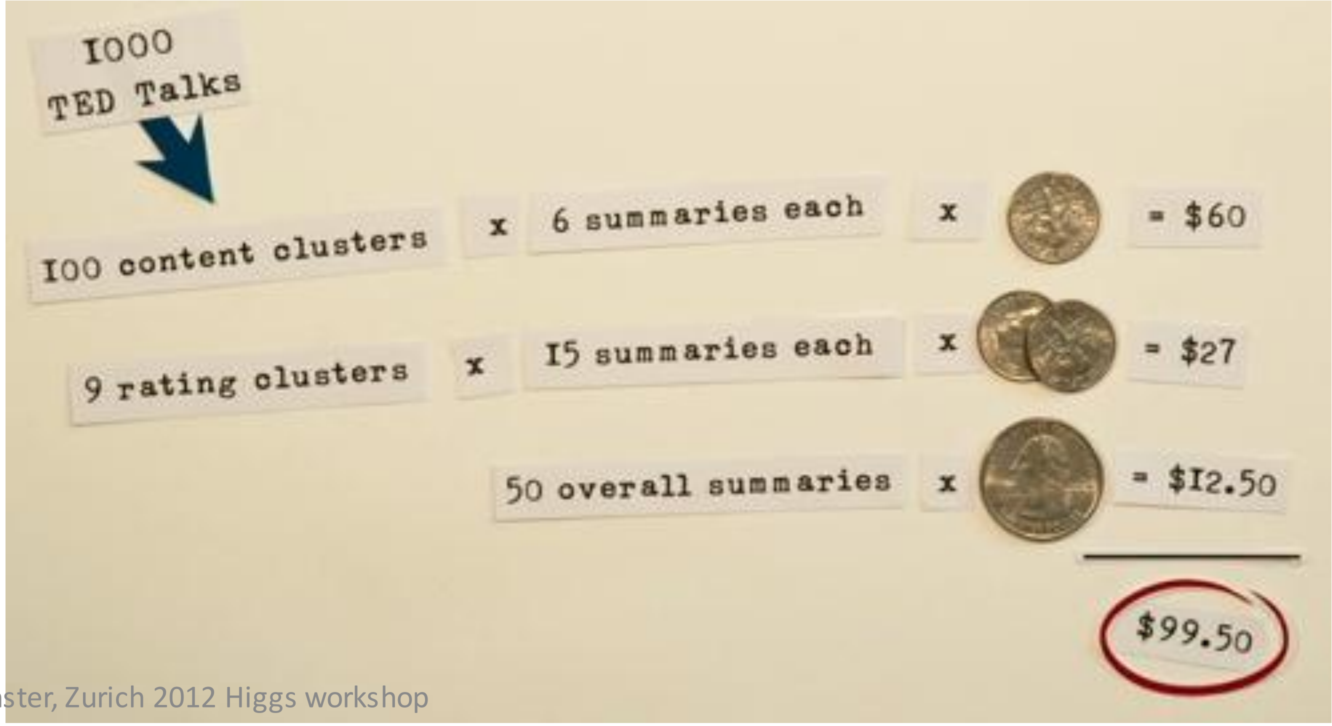
Frederick Balagadde: Bio-lab on a microchip
06:11 Posted: Apr 2010
Views: **102,507** | Comments: **39**
Rated: **Ingenious** Inspiring ...

Nathan Myhrvold: Could this laser zap malaria?
16:58 Posted: May 2010
Views: **292,513** | Comments: **303**
Rated: **Ingenious** Jaw-dropping ...

David Bolinsky animates a cell
09:45 Posted: Jul 2007
Views **671,878** | Comments **178**

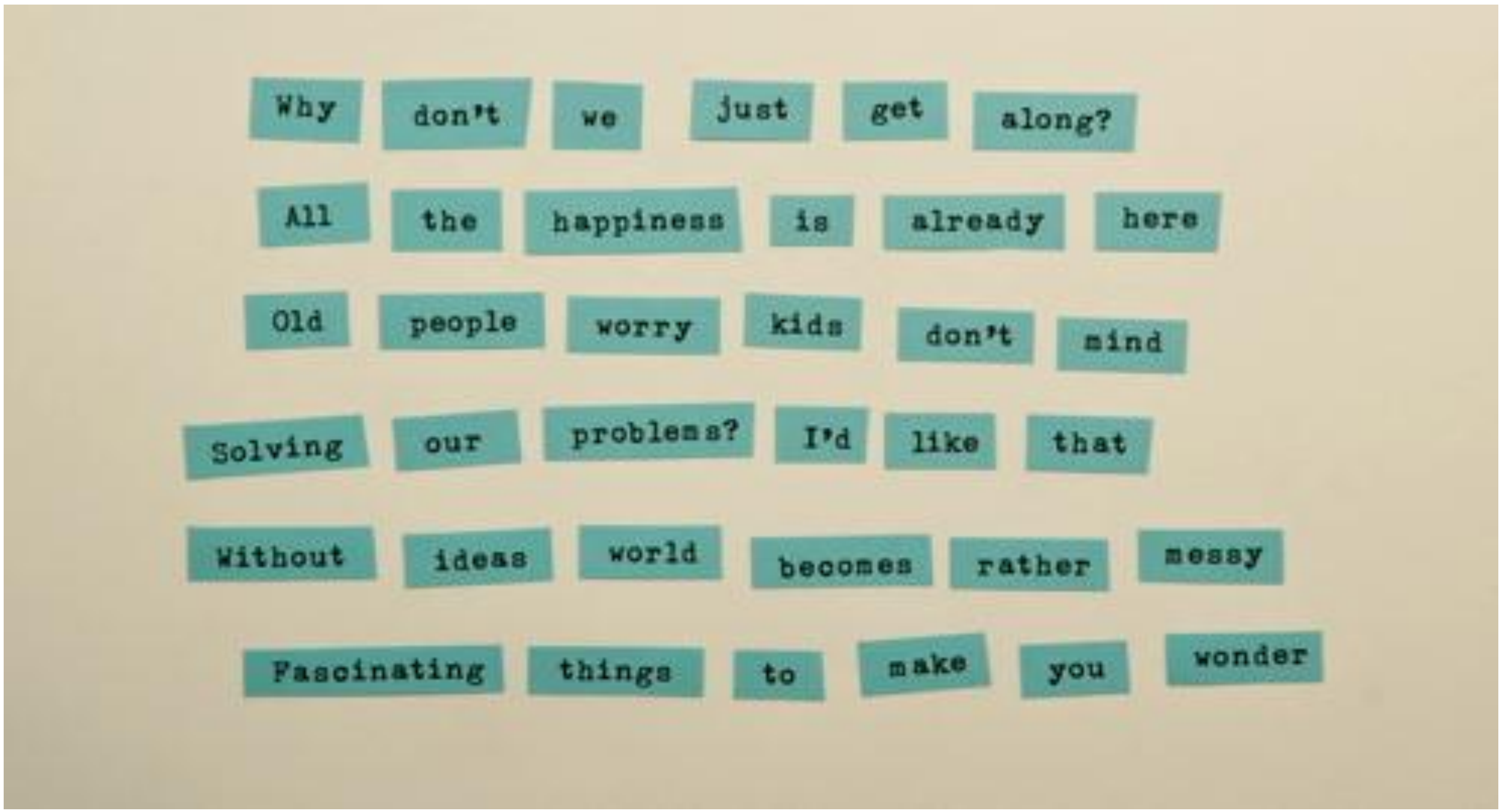
Ted x Zürich, Oct. 2011

- Sebastian Wernicke
  - Leader in field of bioinformatics
  - Used Mechanical Turk website to hire people to do Human Intelligence Tasks for 10 cents each
  - 1000 Ted Talks, each ~2300 words, summarized to 6 words for $100



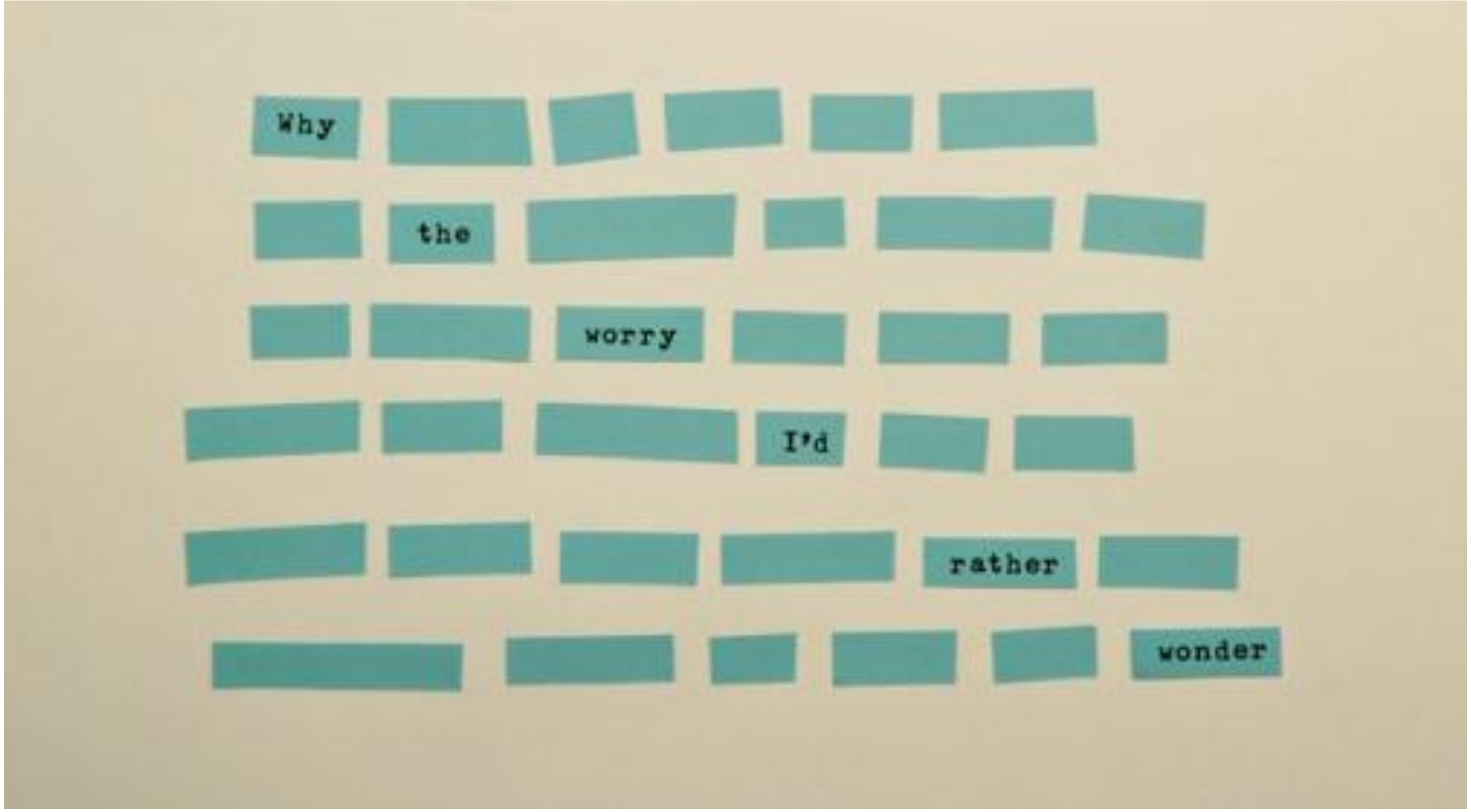
Ted x Zürich, Oct. 2011

Six of the fifty 6-word summaries of all 1000 Ted Talks



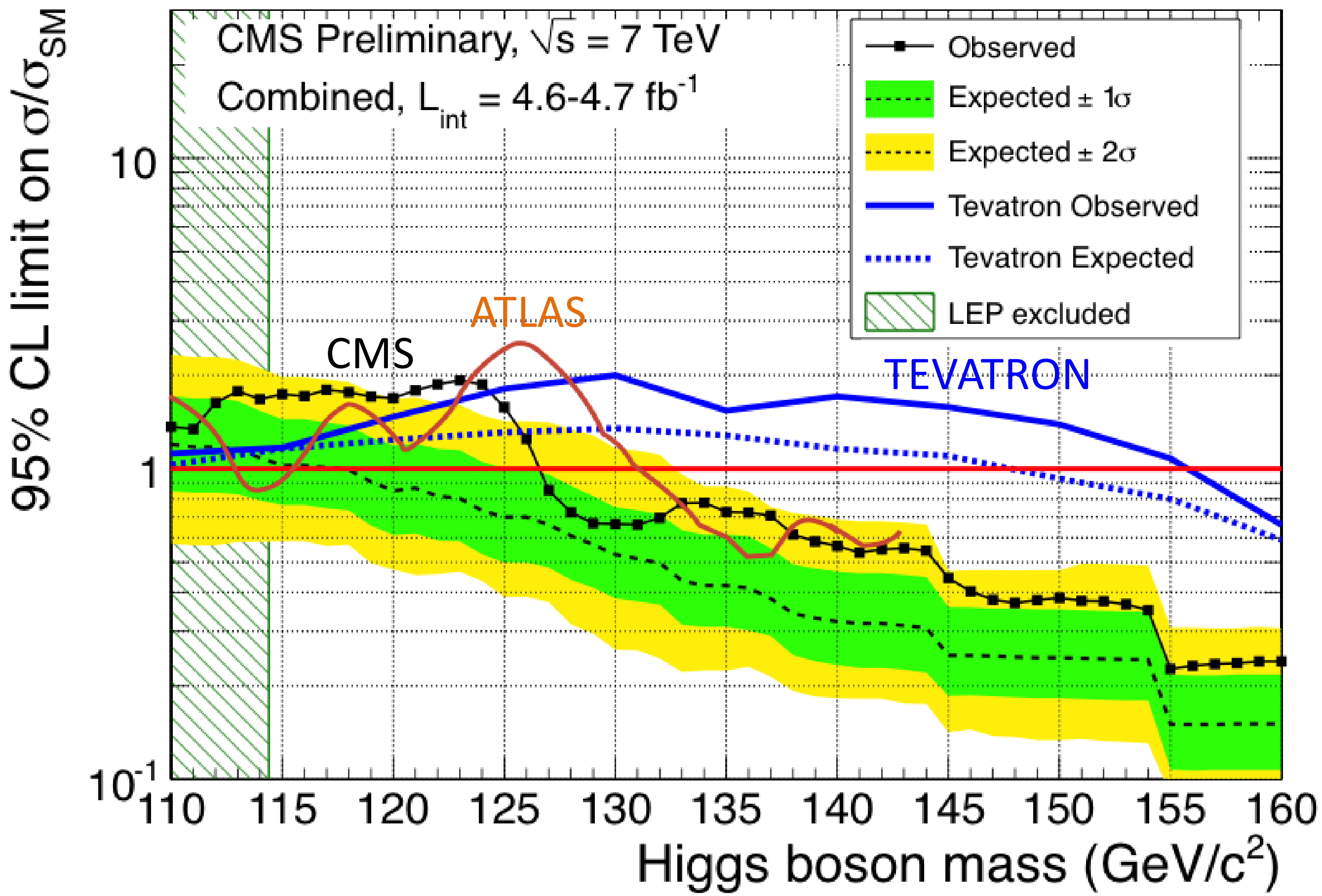
Ted x Zürich, Oct. 2011

Still not satisfied, he chose these 6 final words



Information is clearly lost ☺

Physicists summarize with a picture



Exclusions above this :

Tevatron:
158 – 177 GeV

CMS:
127-600 GeV

ATLAS:
112.7-155.5
131-237
251-468

Caveat : cost >> $100

Except …

- A one-picture summary begs the question :
  - What went into this ?

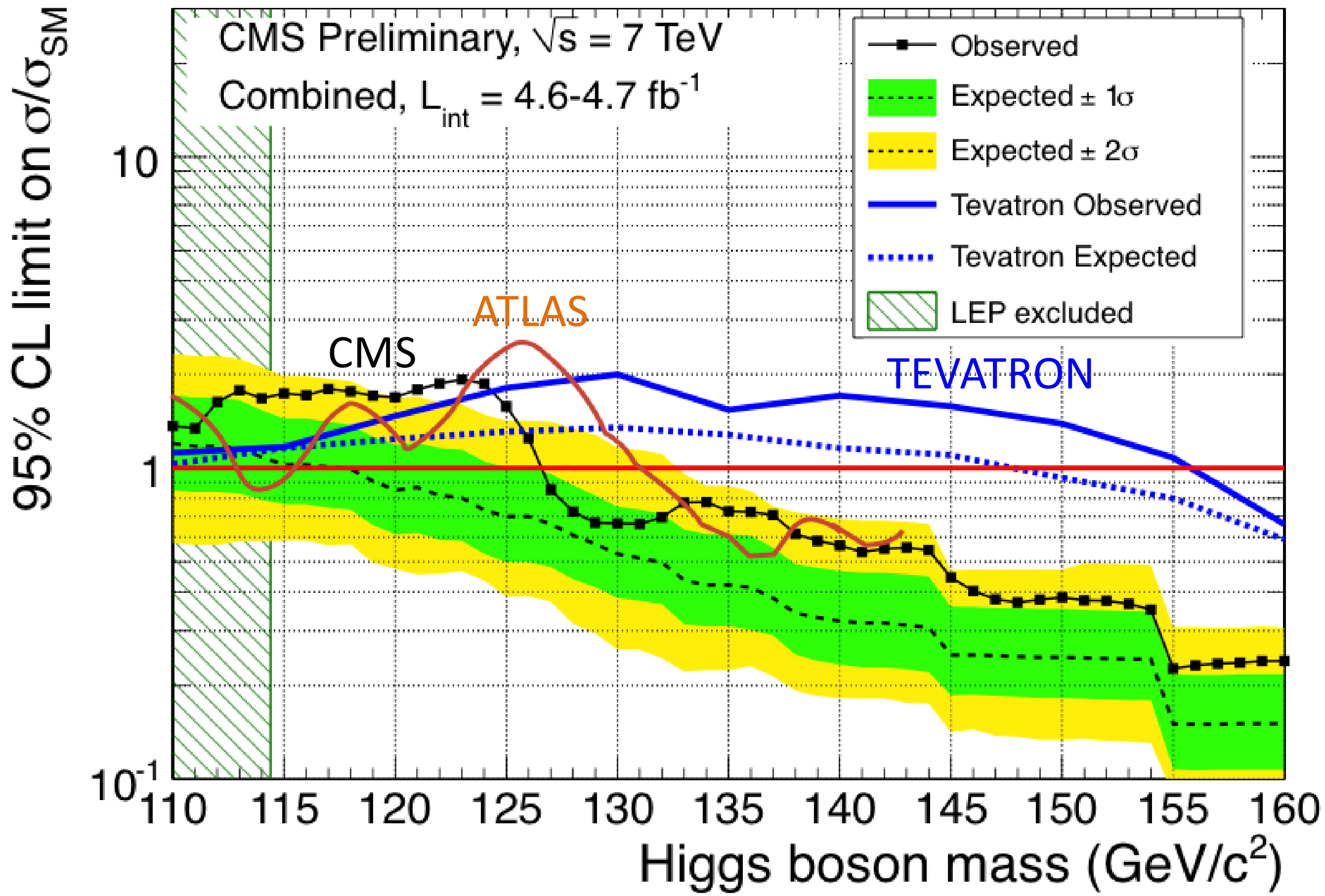
- None of us here accept a single picture as the answer without considering its components

Specifically for our Higgs picture

- What goes into TeV & LHC limits ?
- What are all the pieces ?
- How does it all fit together ?
- What are the assumptions ?
  - How important are they to the conclusion ?
- How do we define excesses and deficits ?
  - What do they tell us ?
- To what degree is it a consistent picture ?

Or in 6 words …

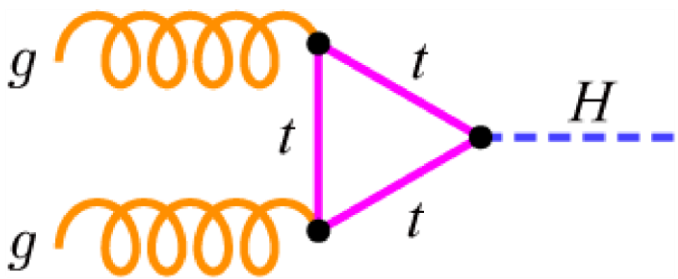
How much information does picture lose ?

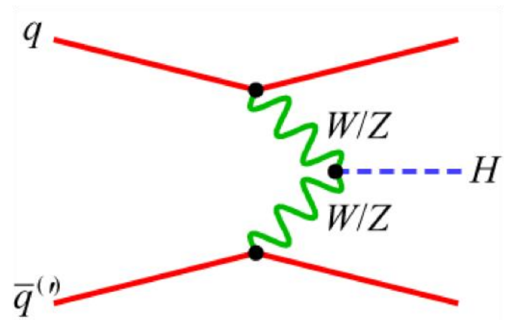


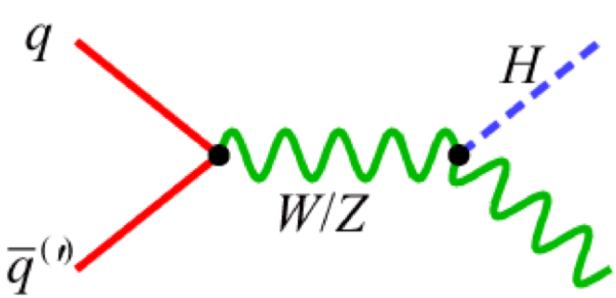
Higgs production

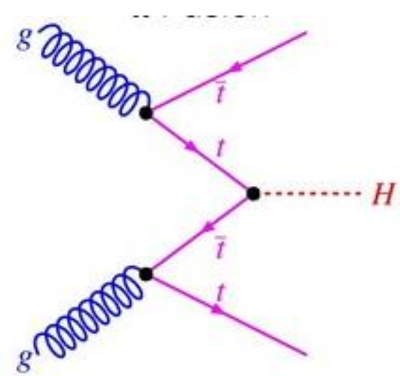
Actually searching for four production modes

- Picture uses single multiplier μ of the SM cross-sections for four different Higgs productions









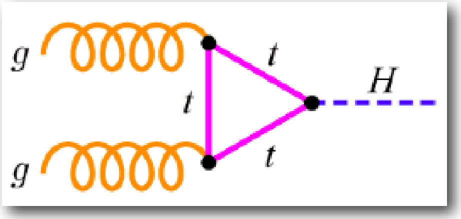
Higgs Production in combinations

Production	LEP	Tevatron	LHC
qq → Z* → ZH			
qq →W* → WH			
gg → H			
qq → WW/ZZ qq → Hqq			
gg → tttt → ttH			

Tevatron vs. LHC

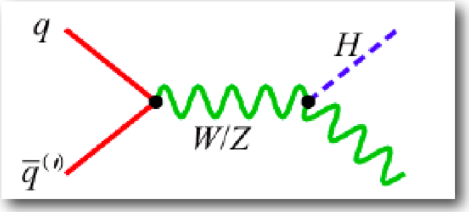
- LHC has higher cross-sections for signal
  - But scaling is not the same
  - Different production cocktail between accelerators

Gluon fusion



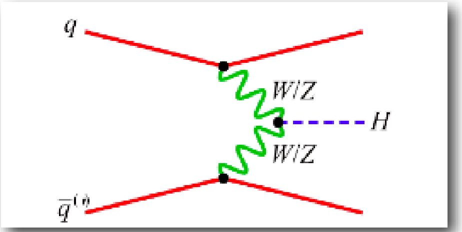
40x more at LHC

Associated Production



10x more at LHC

Vector boson fusion



50x more at LHC

Sensitivity at 115 GeV :

LHC

Tevatron

Higgs production uncertainties

Each Higgs production also comes with different relative uncertainties

At LHC

Source	Affected Processes	Typical uncertainty
PDFs+α_s (cross sections)	$gg \to H$, $t\bar{t}H$, $gg \to VV$	$\pm 8\,\%$
	VBF H, VH, VV@NLO	$\pm 4\,\%$
Higher-order uncertainties on cross sections	total inclusive $gg \to H$	$^{+12\,\%}_{-7\,\%}$
	inclusive "gg" $\to H + \geq 1$ jets	$\pm 20\,\%$
	inclusive "gg" $\to H + \geq 2$ jets	$\pm 20\%$ (NLO), $\pm 70\%$ (LO)
	VBF H	$\pm 1\,\%$
	associated VH	$\pm 1\,\%$
	$t\bar{t}H$	$^{+4\,\%}_{-10\,\%}$
	uncertainties specific to high mass Higgs boson, see Section 2.1	$\pm 30\%$

Correlated between all channels and each experiment

gg → H uncertainties are largest despite tremendous set of calculations :

- QCD radiative corrections at NLO
- QCD corrections NNLO
- QCD soft-gluon resummation NNLL
- EWK corrections NLO
- top and bottom loop corrections up NLO
- above 400 GeV, line shape unknown

Details & references in CMS+ATLAS combination note

Higgs exclusive production uncertainties

Production modes have exclusive uncertainties

- Splitting H→WW by number of jets
  - Different PDF+α_s and scale errors for each jet-bin
    - PDF errors from Anastasiou et al,. JHEP 0908, 099 (2009)
  - Treat scale uncertainty of NNLO+NNLL inclusive, but NLO 1+ jet, 2+jet bins as uncorrelated
    - Berger et al., arXiv:1012.4480, Stewart and Tackman, arXiv:1107:2217

At Tevatron

- 3 scales - Tackmann et al., arXiv:1107.2217 [hep-ph] → 3 nuisance parameters
  - S0 - scale uncertainty on x0, S1 - scale uncertainty on x1, S2 - scale uncertainty on x2

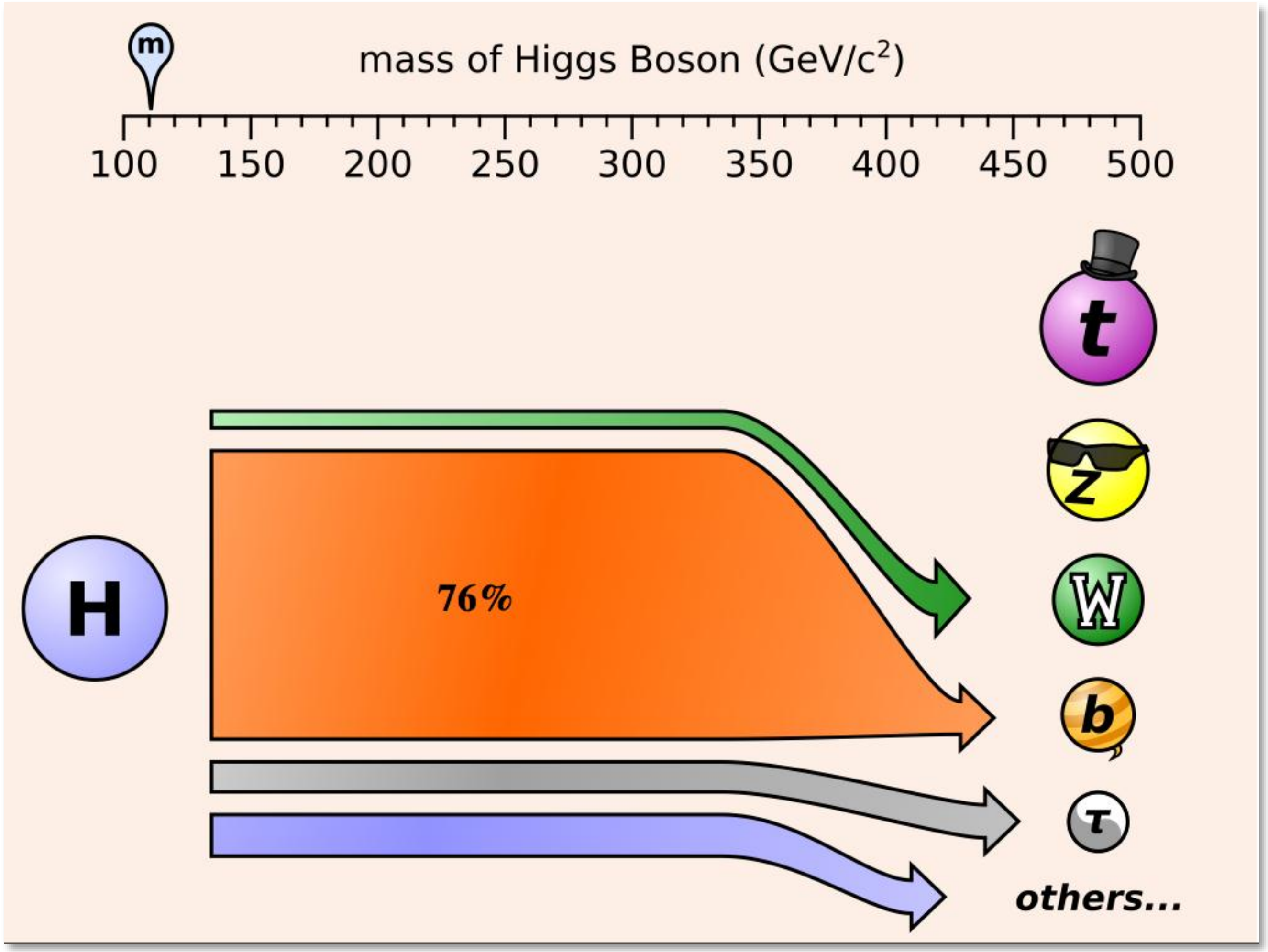
- X0: Inclusive cross section: Florian & Grazzini, Phys. Lett. B 674, 291 (2009)
- X1: H+1-or-more-jets: MCFM
- x2: H+2-or-more-jets: Campbell, Ellis & Williams, arXiv:1001.4495 [hep-ph]

Signal Category	S0	S1	S2
0-jet	S0x(x0/(x0-x1))	-S1x(x1/(x0-x1))	0
1-jet	0	S1x(x1/(x1-x2))	-S2x(x2/x1-x2))
2-jet	0	0	S2

Jet bin	s0	s1	s2
0 jet	13.4%	-23.0%	0
1 jet	0	35%	-12.7%
>= 2 jets	0	0	33%

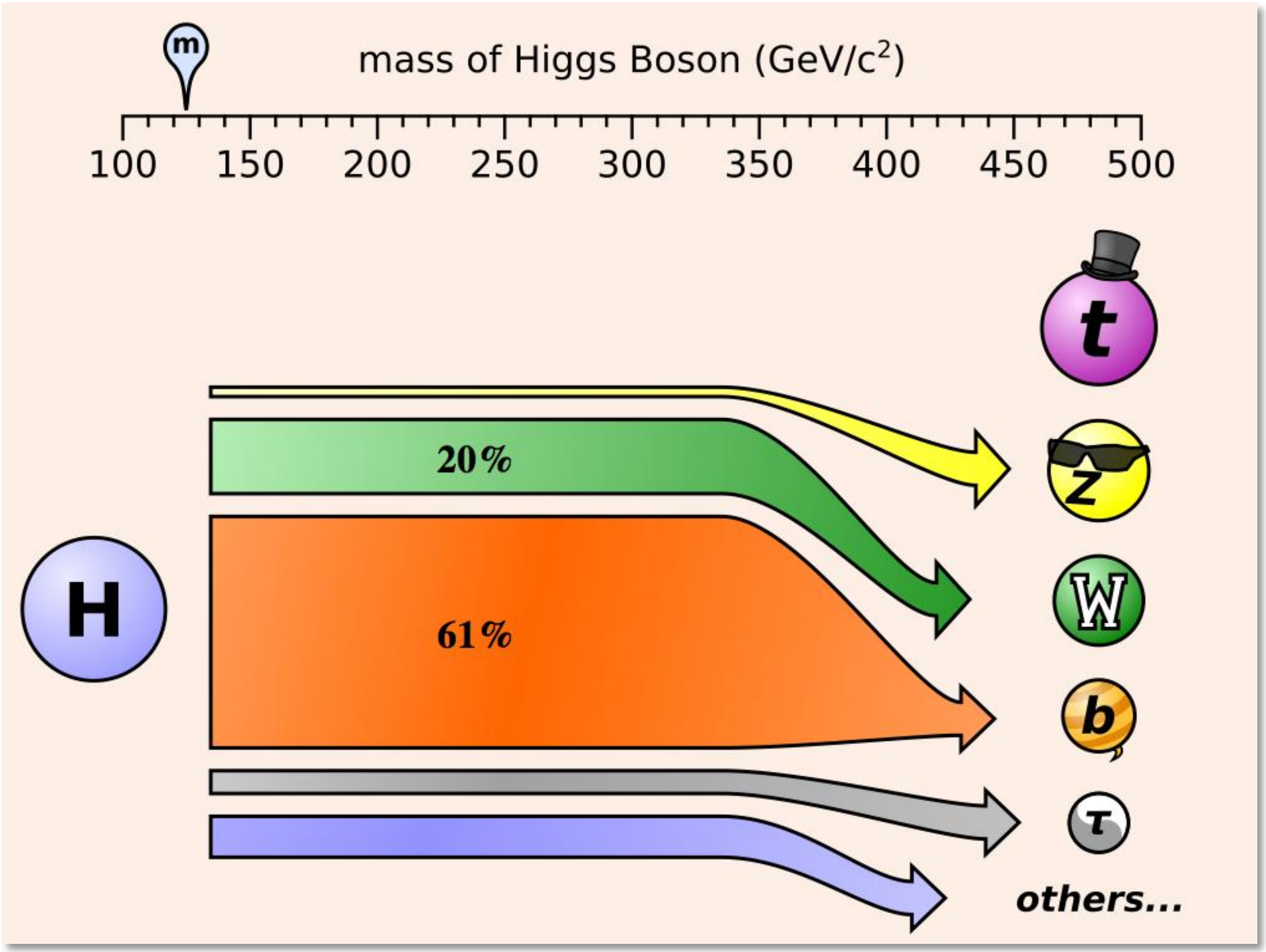
Higgs decays

110 GeV



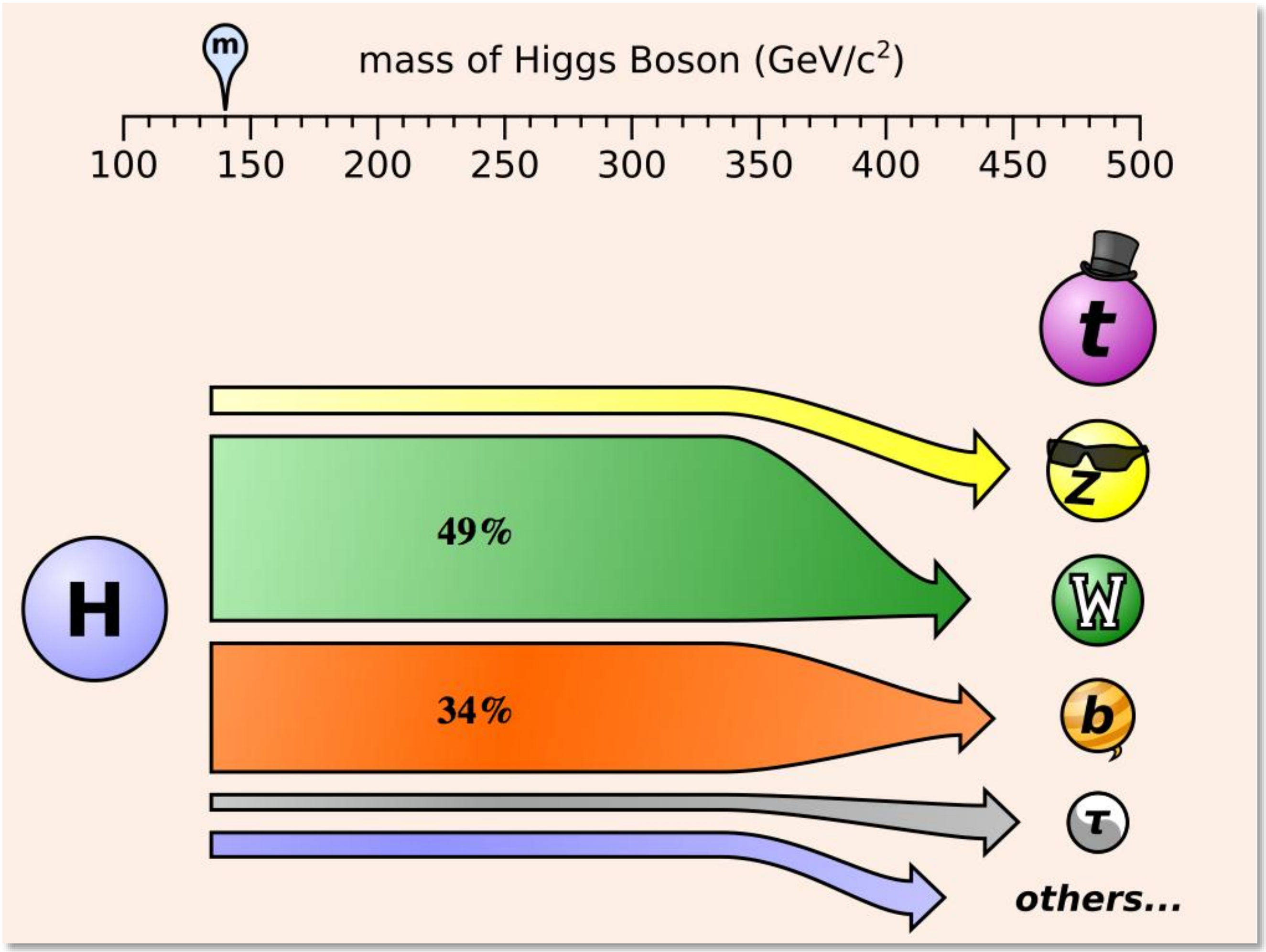
Comes from svg animation by Jim Pivasrksi

The infamous 125 GeV excess

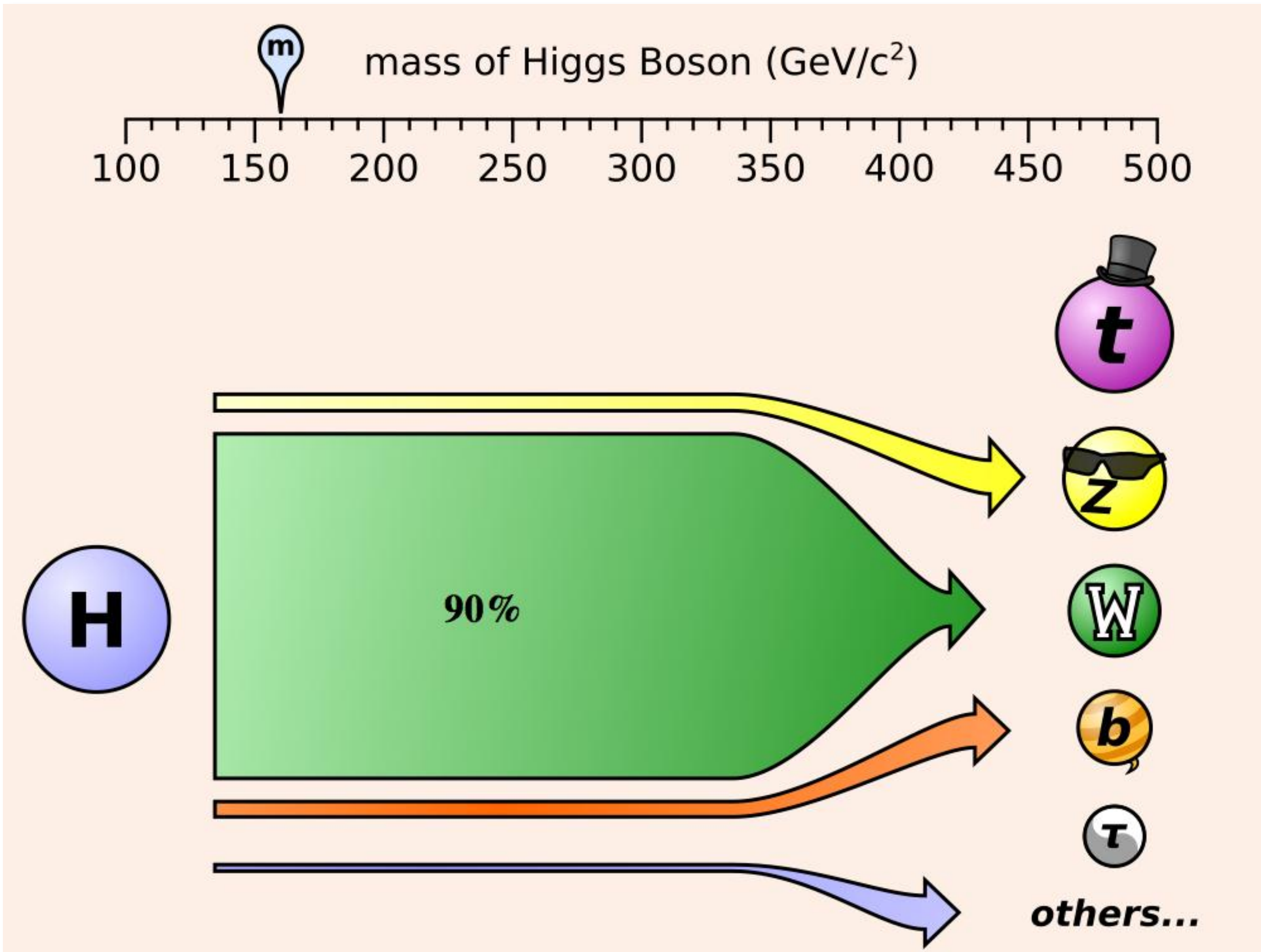


NOTE:
Excess not due to dominant Higgs decay

140 GeV

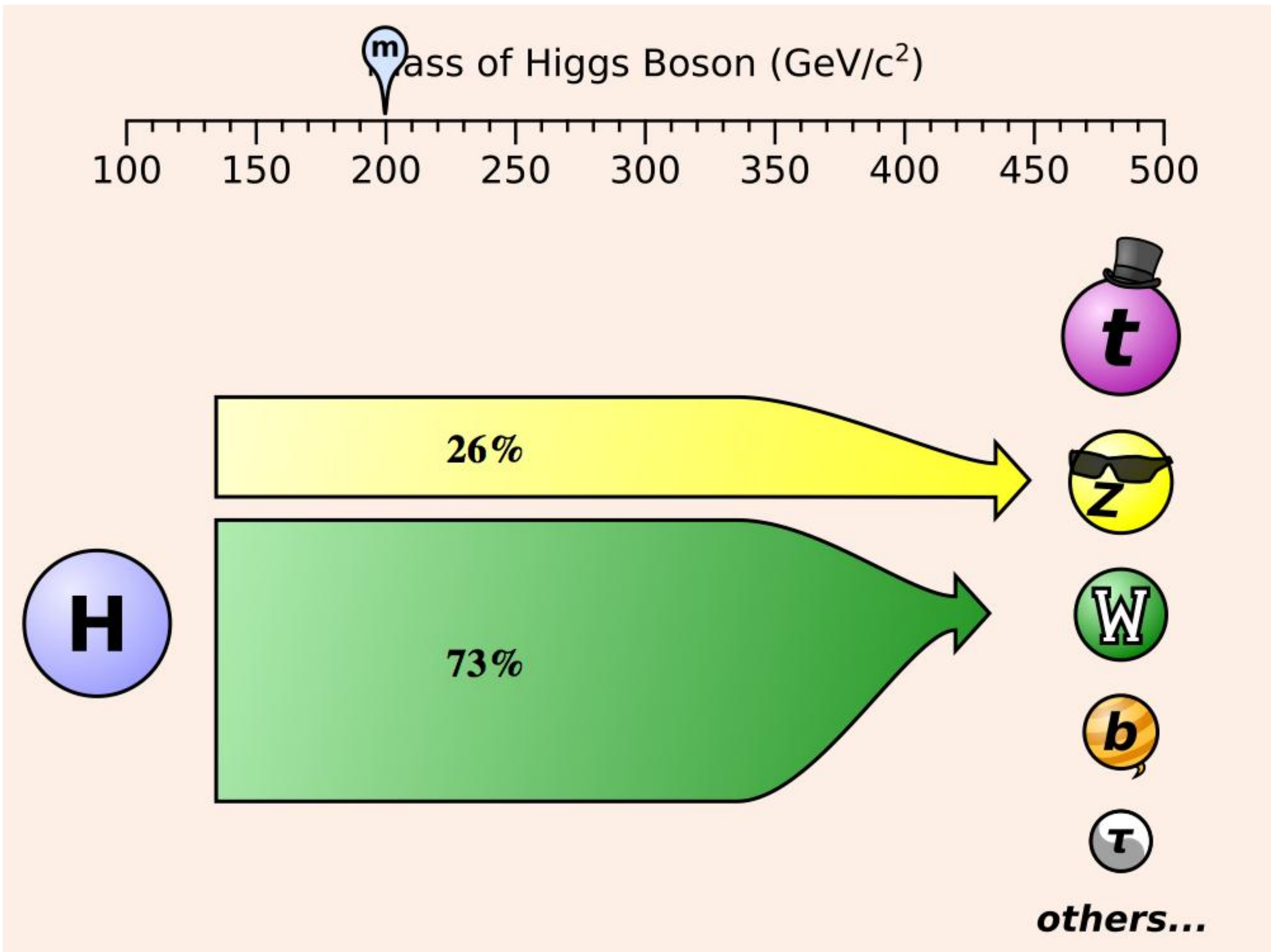

mass of Higgs Boson (GeV/c²)
m
100
150
200
250
300
350
400
450
500
H
49%
34%
t
Z
W
b
τ
others...

160 GeV



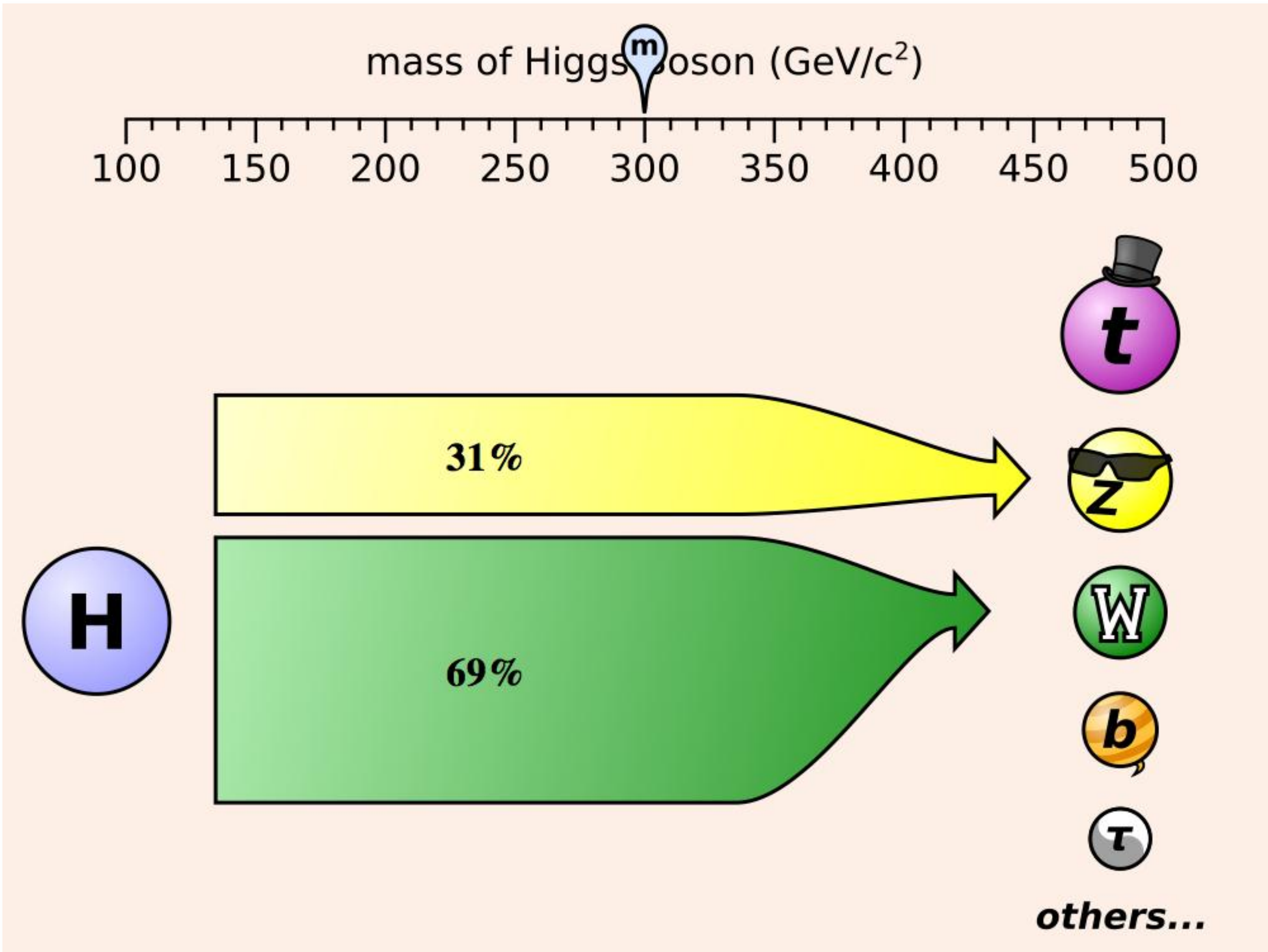
m
mass of Higgs Boson (GeV/c²)
100
150
200
250
300
350
400
450
500
H
90%
t
Z
W
b
τ
others...

200 GeV



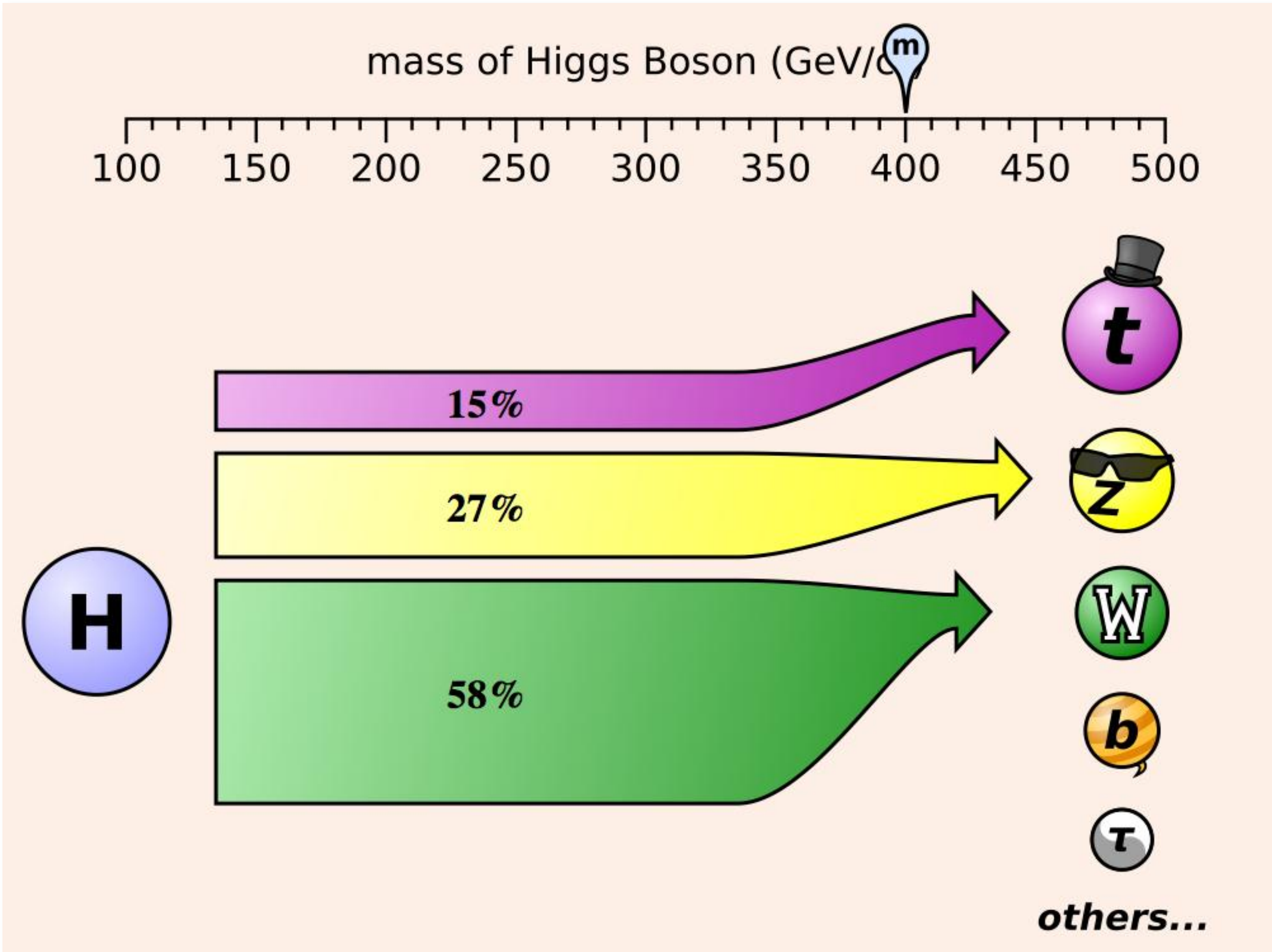
Mass of Higgs Boson (GeV/c²)
100 150 200 250 300 350 400 450 500
26%
73%
H
t
Z
W
b
τ
others...

300 GeV



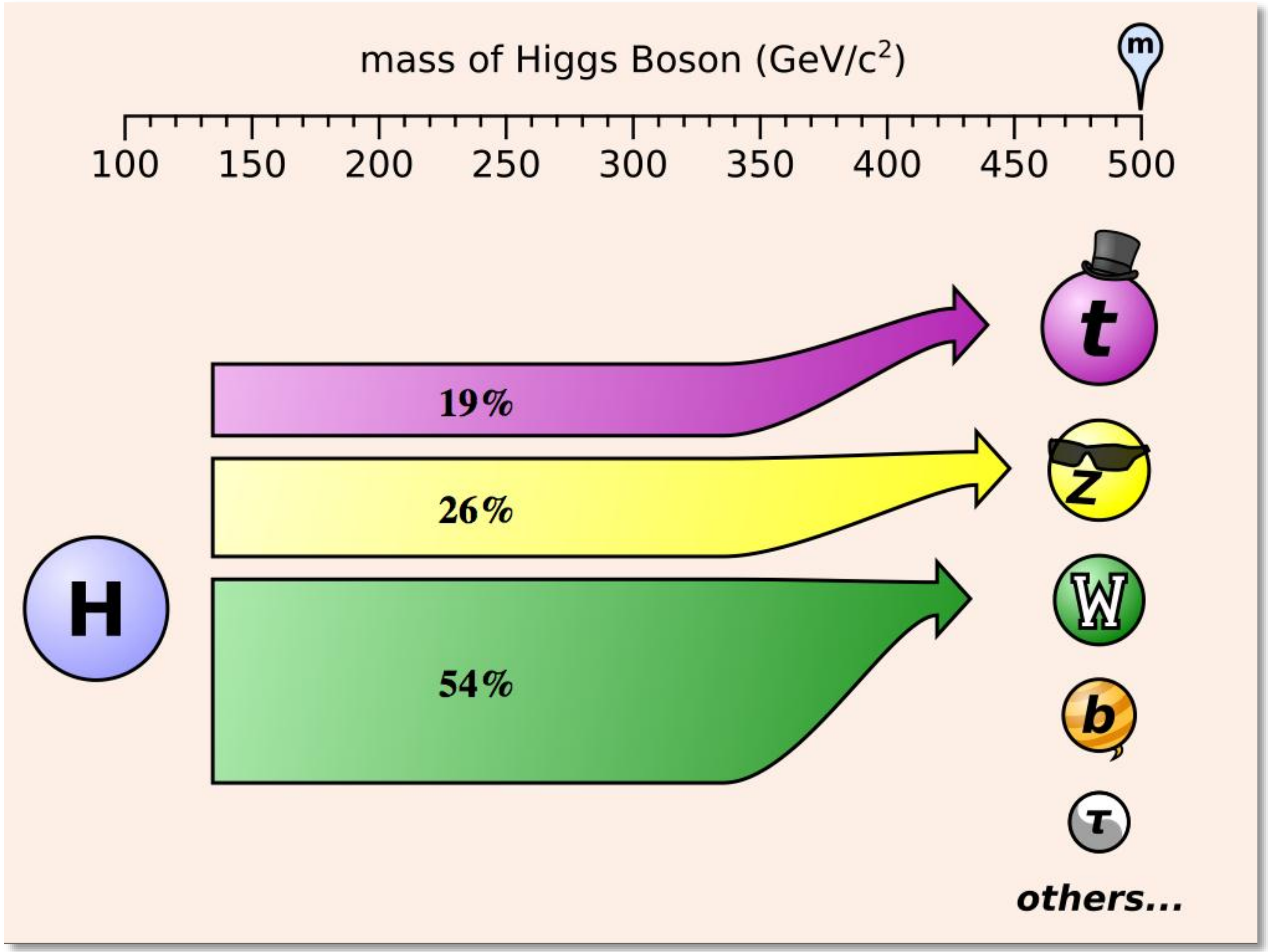
mass of Higgs boson (GeV/c²)
m
100
150
200
250
300
350
400
450
500
31%
69%
H
t
Z
W
b
τ
others...

400 GeV



mass of Higgs Boson (GeV/c
m
100
150
200
250
300
350
400
450
500
15%
27%
58%
H
t
Z
W
b
τ
others...

500 GeV



NOTE: Exclusion at 525 GeV does not consider H→tt

Channels in picture

Associated Higgs modes

	LEP	Tevatron	LHC
ZH→vv(bb)			
ZH→qq(bb)			
ZH→ll(bb)			
ZH→ττ(bb)			
ZH→qq(ττ)			
ZH→ZWW→lll			
WH→lv(bb)			
WH→qq(bb)			
WH→τv(bb)			
WH→WWW→ll(l)			

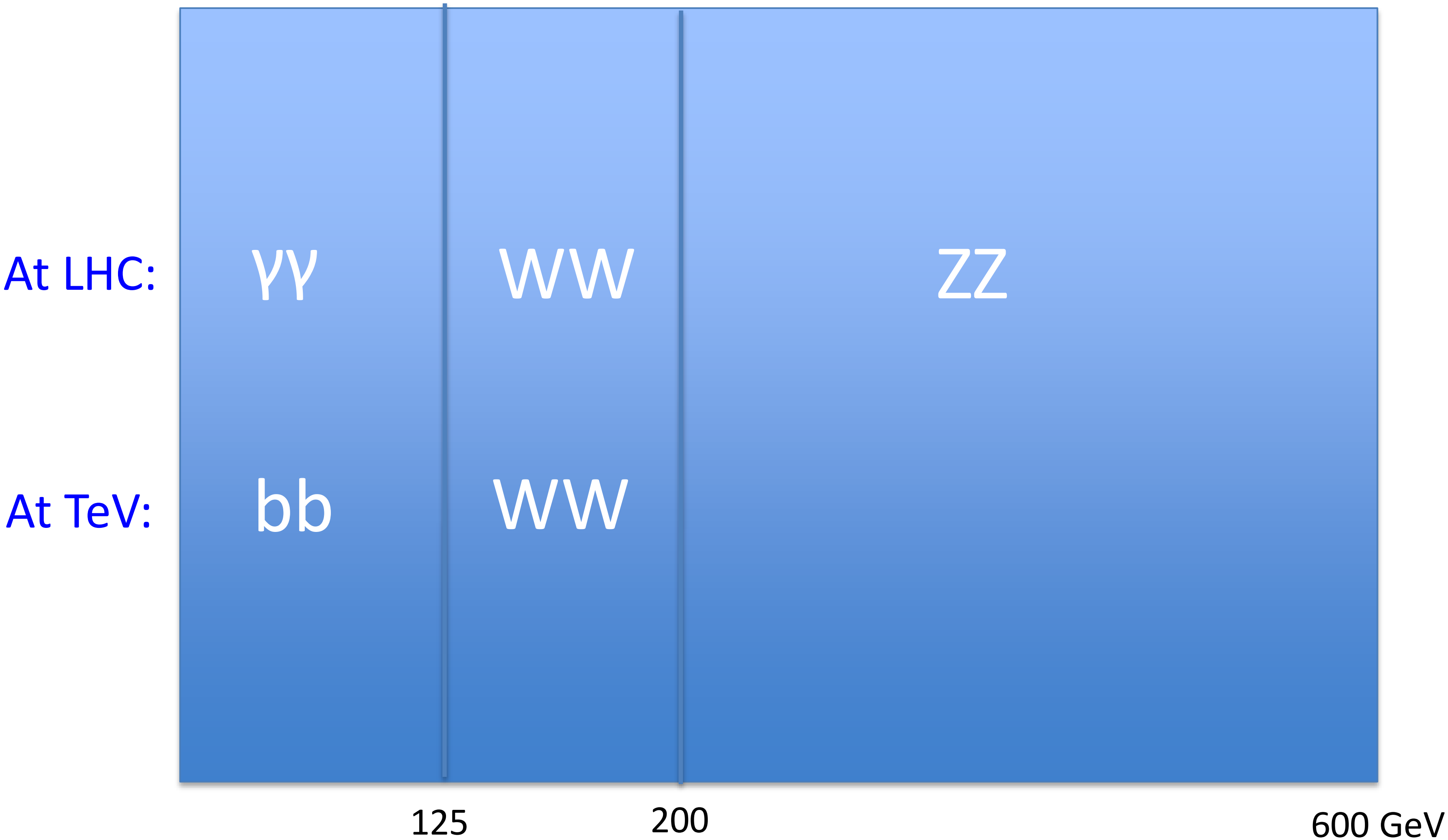
Channels in picture

Gluon fusion, VBF, ttH

	Tevatron	LHC
H→WW→lvlv		
H→WW→lvqq		
H→WW→lvτv		
H→ZZ→llll		
H→ZZ→llvv		
H→ZZ→llqq		
H→ZZ→llττ		
H→ZZ→vvqq		
H→ττ+ jets		
H→γγ		
ttH → lv+bb(b)		
ttH → MET+bb(b)		
ttH → qq+bb(b)		

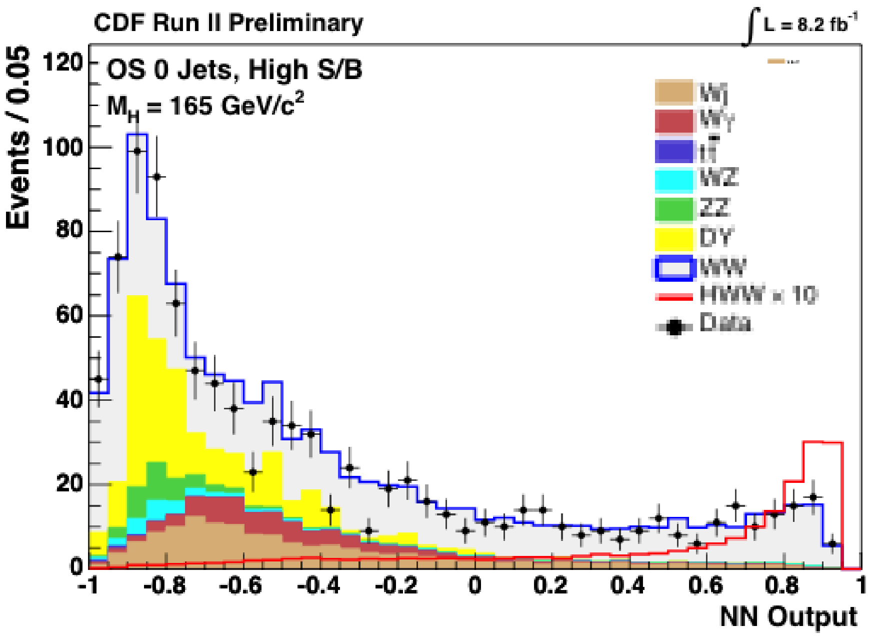
Higgs backgrounds

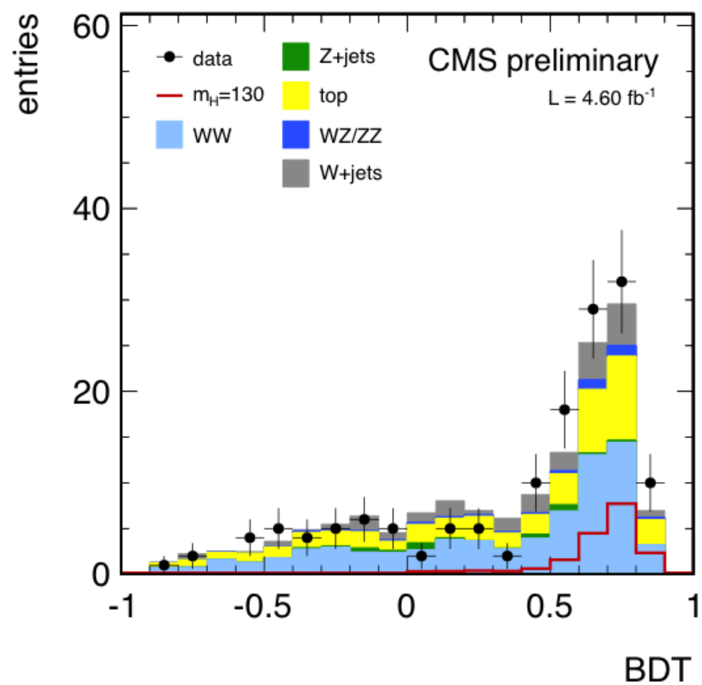
Different dominant SM backgrounds at each mass



Higgs background composition at LHC vs. Tevatron

WW searches with 0 jets





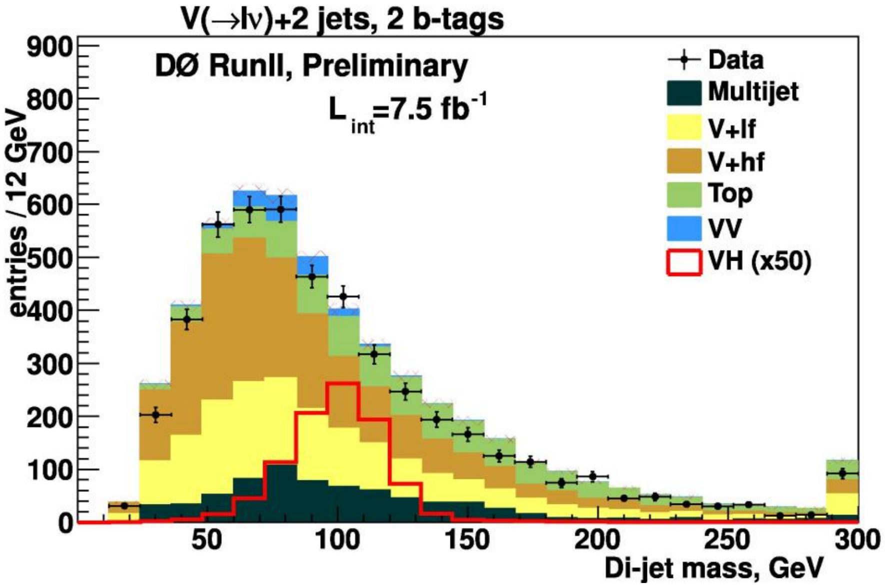
Tevatron: larger Drell-Yan

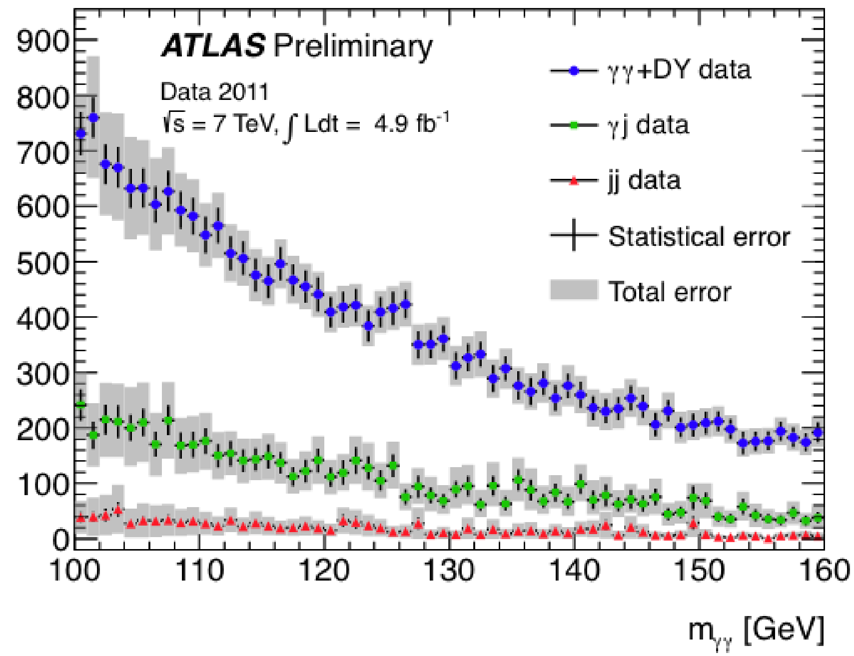
LHC : larger tt

Consistent limits between Tevatron and LHC make background mis-modeling less likely

Higgs background composition at LHC vs. Tevatron

Low mass range



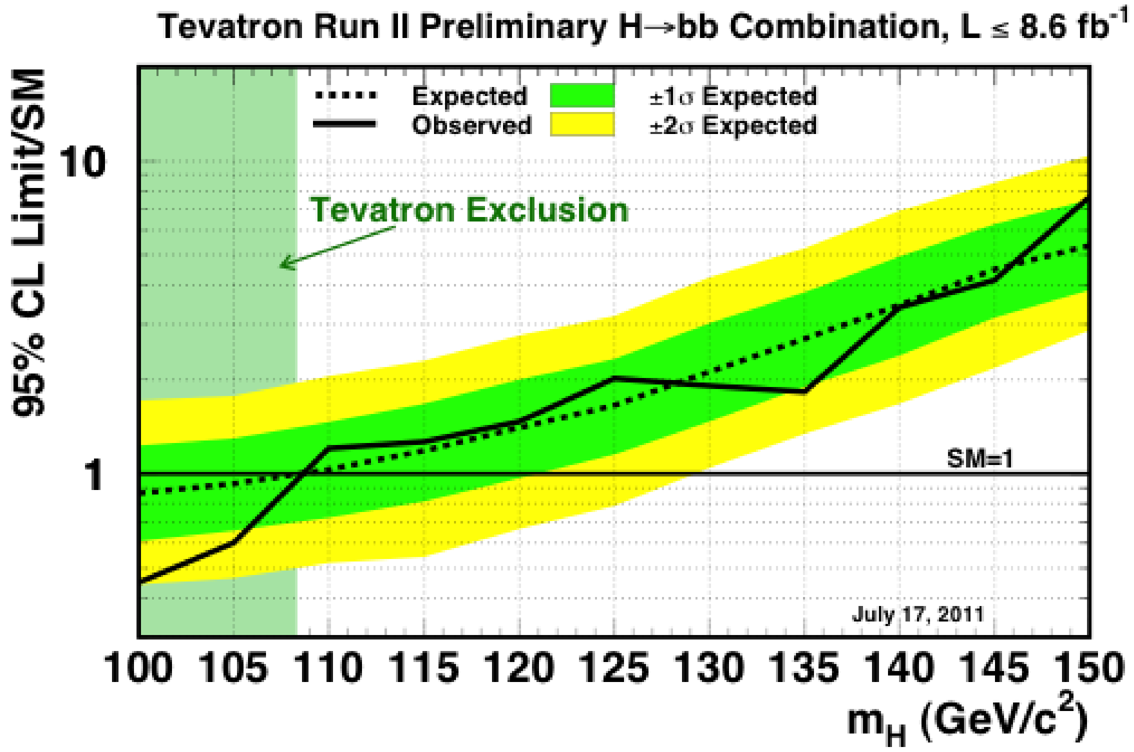


Tevatron: W+jets dominant in H→bb

LHC : γγ dominant in H→γγ

Consistent excesses between Tevatron and LHC would make background mis-modeling less likely

Tevatron H→bb



NOTE: H→bb excess has not developed but, if there, should be expected across this mass range with full dataset

Statistical techniques

Different test statistics used in picture

Table 11: Comparison of CL_s definitions as used at LEP, Tevatron, and adopted for the summer 2011 Higgs combination at LHC.

	Test statistic	Profiled?	Test statistic sampling
LEP	$q_\mu = -2 \ln \frac{\mathcal{L}(data\|\mu,\tilde{\theta})}{\mathcal{L}(data\|0,\tilde{\theta})}$	no	Bayesian-frequentist hybrid
Tevatron *	$q_\mu = -2 \ln \frac{\mathcal{L}(data\|\mu,\hat{\theta}_\mu)}{\mathcal{L}(data\|0,\hat{\theta}_0)}$	yes	Bayesian-frequentist hybrid
LHC	$\tilde{q}_\mu = -2 \ln \frac{\mathcal{L}(data\|\mu,\hat{\theta}_\mu)}{\mathcal{L}(data\|\hat{\mu},\hat{\theta})}$	yes $(0 \le \hat{\mu} \le \mu)$	frequentist

* Tevatron quotes limits with Bayesian technique; CL_S is a cross-check

From CMS+ATLAS combination procedure note

μ : scaling of signal cross-section where SM=1
θ : nuisance parameters
q_μ : test statistic of the signal + background model

CL_s technique

- Confidence levels are evaluated by integrating corresponding log likelihood ratio distributions populated by simulating outcomes via Poisson statistics
- LHC: Pseudo-data is generated using best fit of nuisance parameters to the observed data
  - For both background-only and signal+background hypothesis in LLR
- Tevatron: Pseudo-data is generated using expected values of nuisance parameters

- CL_s is computationally expensive
  - LHC CL_s has asymptotic properties so that limits can be evaluated with a simple formula – no pseudo-data needed :

$$CL_s = 0.05 = \frac{1 - \Phi(\sqrt{q_\mu})}{\Phi(\sqrt{q_{\mu,A}} - \sqrt{q_\mu})}$$

Φ^{-1} is the quantile (inverse of the cumulative distribution) of the standard Gaussian.

Bayesian technique used by Tevatron

Bayesian Posterior Probability

$$p(R|\vec{n}) = \frac{\int\int d\vec{s}d\vec{b}L(R,\vec{s},\vec{b}|\vec{n})\pi(R,\vec{s},\vec{b})}{\int\int\int dRd\vec{s}d\vec{b}L(R,\vec{s},\vec{b}|\vec{n})\pi(R,\vec{s},\vec{b})} \Rightarrow \int_0^{R_{0.95}} p(R|\vec{n})dR = 0.95$$

$R = (\sigma \times BR)/(\sigma_{SM} \times BR_{SM})$, $R_{0.95}$: 95% Credible Level Upper Limit

$\vec{s}, \vec{b}, \vec{n} = s_{ij}, b_{ij}, n_{ij}$ (# of signal, background and observed events in j-th bin for i-th channel)

π : Bayes' prior density

Combined Binned Poisson Likelihood

$$L(R,\vec{s},\vec{b}|\vec{n}) = \prod_{i=1}^{N_{\text{channel}}} \prod_{j=1}^{N_{\text{bin}}} \frac{\mu_{ij}^{n_{ij}} e^{-\mu_{ij}}}{n_{ij}!}$$

Principle of ignorance

- for the number of higgs events (instead of higgs Xsec)

$$\pi(R,\vec{s},\vec{b}) = \pi(R)\pi(\vec{s})\pi(\vec{b}) = s_{tot}\theta(Rs_{tot})\pi(\vec{s})\pi(\vec{b})$$

$s_{tot} = \Sigma_{i,j} s_{ij}$: Total number of signal prediction

$\pi(x) = G(x|\hat{x}, \sigma_x)$ $(x = s, b)$ $\hat{x}$: expected mean, σ_x: total uncertainty

The integrals over the uncertain parameters with their correlated priors from external constraints are done with a Markov Chain Monte Carlo integration method, using the Metropolis-Hastings algorithm.

- CLs vs Bayesian ?
- Different flavors of CL_s (LEP, Tevatron, LHC)
- Asymptotic approximation of CL_s without toys

Do we need to care what is used ?

Not obvious from get-go, but the answer is “NO”

CL_s vs. Bayesian

- Tevatron limits from summer 2011

TABLE V: Ratios of median expected and observed 95% C.L. limit to the SM cross section for the combined CDF and D0 analyses as a function of the Higgs boson mass in GeV/c^2, obtained with the Bayesian and with the CL_s method.

Bayesian	155	160	165	170	175	180	185	190	195	200
Expected	0.80	0.59	0.57	0.67	0.80	0.97	1.22	1.49	1.71	2.02
Observed	1.08	0.66	0.48	0.62	0.91	1.14	1.31	1.90	2.41	2.91

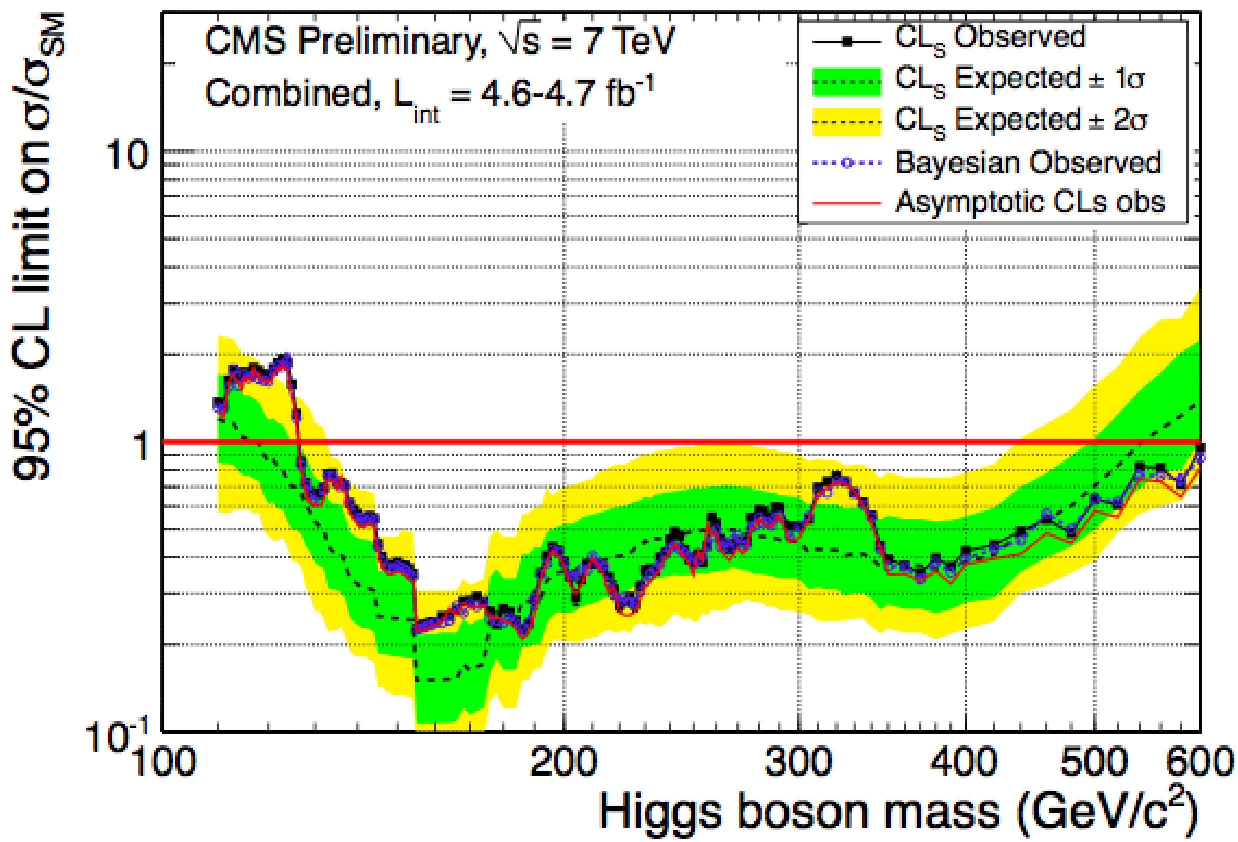
CL_s	155	160	165	170	175	180	185	190	195	200
Expected	0.82	0.61	0.58	0.67	0.81	0.98	1.24	1.50	1.77	2.04
Observed	1.03	0.67	0.48	0.61	0.92	1.17	1.34	1.92	2.39	2.82

- Expected agree to 1-2% on average
- Observed agree to 1-3 % or so on average
- Max disagreement is 2.23 -> 2.38 (10%)

Two philosophies draw same conclusions

Asymptotic vs. CL_s vs. Bayesian

- CMS Dec. 2011 combination

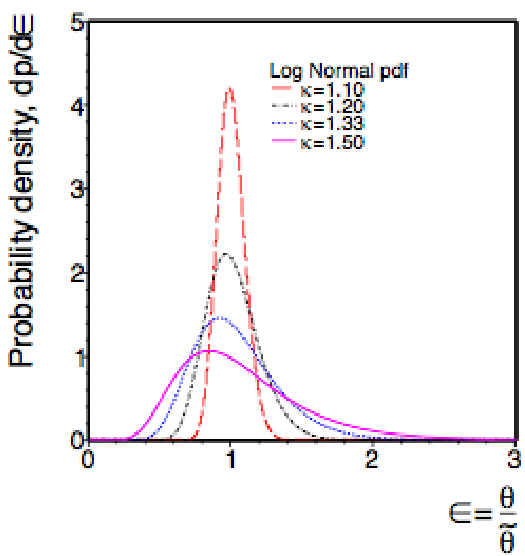


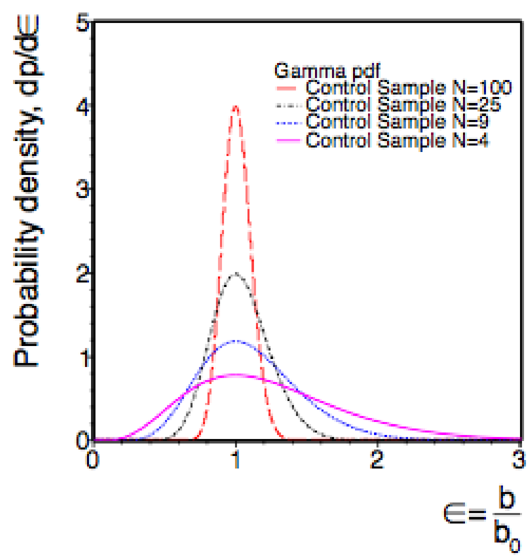
Strong agreement – asymptotic agrees better when high statistics

Treatment of nuisance parameters

Choice of PDFs for nuisance parameters

- Flat or uniform priors
  - Ie, constrained by data measurement, such as signal cross-section
- Poisson
  - Ie, constrained from event counts in control regions or MC statistics
- Normal
  - Gaussian
  - If uncertainties can assume only positive values
    - Log-normal - LHC
    - Truncated - Tevatron (less elegant, but found to be same as log-normal)





For small uncertainties, or large statistics, log-normal and Gamma distribution equivalent to Gaussian

Correlated between analyses and experiments

PDF+α_s uncertainties

nuisance	groups of physics processes
pdf_gg	$gg \to H$, $t\bar{t}H$, VQQ, $t\bar{t}$, tW, tb (s-channel), $gg \to VV$
pdf_qqbar	VBF H, VH, V, VV, $\gamma\gamma$
pdf_qg	tbq (t-channel), γ+jets

QCD scale uncertainties

nuisance	groups of physics processes
QCDscale_ggH	total inclusive $gg \to H$
QCDscale_ggH1in	inclusive $gg/qg \to H+ \geq 1$ jets
QCDscale_ggH2in	inclusive $gg/qg \to H+ \geq 2$ jets
QCDscale_qqH	VBF H
QCDscale_VH	associate VH
QCDscale_ttH	$t\bar{t}H$
QCDscale_V	W and Z
QCDscale_VV	WW, WZ, and ZZ up to NLO
QCDscale_ggVV	$gg \to WW$ and $gg \to ZZ$
QCDscale_ZQQ	Z with heavy flavor $q\bar{q}$-pair
QCDscale_WQQ	W with heavy flavor $q\bar{q}$-pair
QCDscale_ttbar	$t\bar{t}$, single top productions are lumped here for simplicity

Phenomenological uncertainties

nuisance	groups of physics processes
UEPS	all processes sensitive to modeling of UE and PS

Acceptance uncertainties

nuisance	comments
QCDscale_WW_EXTRAP	extrap. factor α for deriving WW bkgd in HWW analysis
QCDscale_ttbar_EXTRAP	extrap. factor α for deriving $t\bar{t}$ bkgd in HWW analysis

Instrumental uncertainties

nuisance	comments
lumi	uncertainties in luminosities

Instrumental uncertainties not correlated between experiments

Sometimes correlated between analyses within an experiment depending on measurement technique

Similar for Tevatron

What if there is an excess ?

To quantify an excess of events, we use the alternative test statistic q_0, defined as follows:

$$q_0 = -2\ln\frac{\mathcal{L}(\text{data}\,|\,0,\hat{\theta}_0)}{\mathcal{L}(\text{data}\,|\,\hat{\mu},\hat{\theta})} \quad \text{and } \hat{\mu} \geq 0. \tag{6}$$

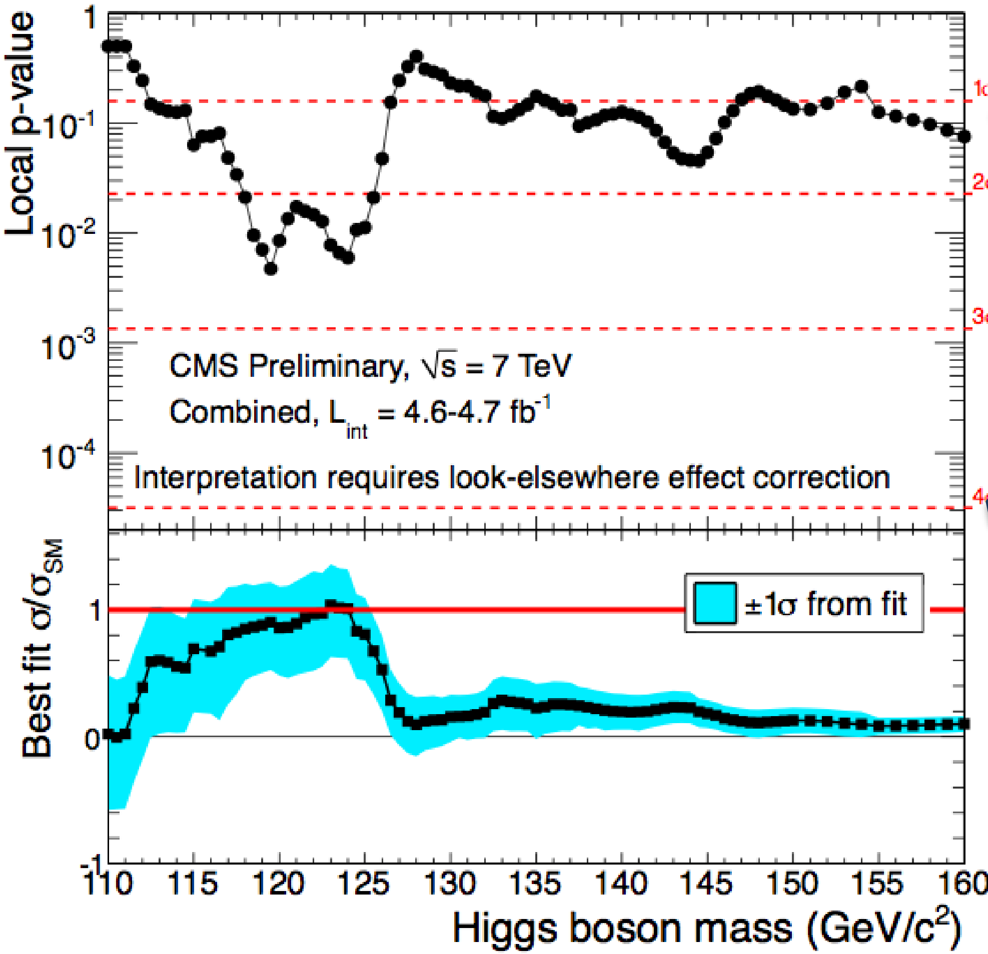
This test statistic allows us to evaluate significances (Z) and p-values (p_0) from the following asymptotic formula [24]:

$$Z = \sqrt{q_0^{\text{obs}}}, \tag{7}$$

$$p_0 = P(q_0 \geq q_0^{\text{obs}}) = \frac{1}{2}\left[1 - \text{erf}\left(Z/\sqrt{2}\right)\right], \tag{8}$$

where q_0^{obs} is the observed test statistic calculated for $\mu = 0$ and with only one constraint $0 \leq \hat{\mu}$, which ensures that data deficits are not counted on an equal footing with data excesses.

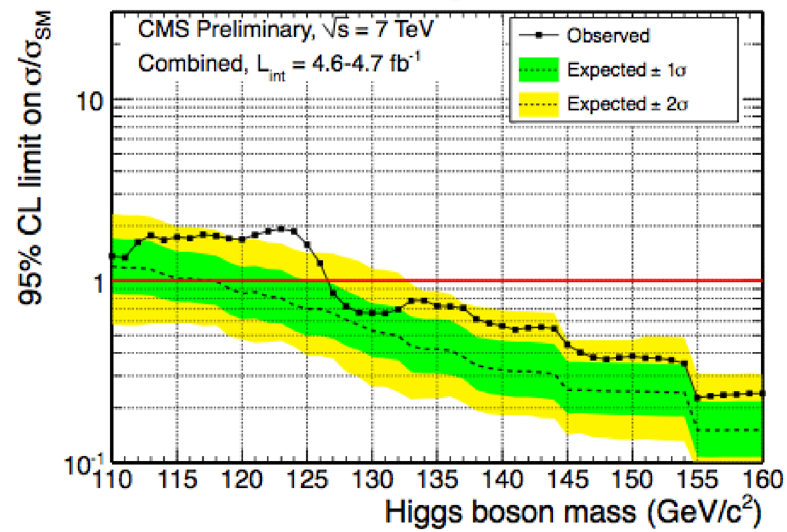
Need complete picture to understand excess



Tells us where the excesses are respect to background

Tells us if excess is consistent with signal

Tells us if we **should** have sensitivity to it

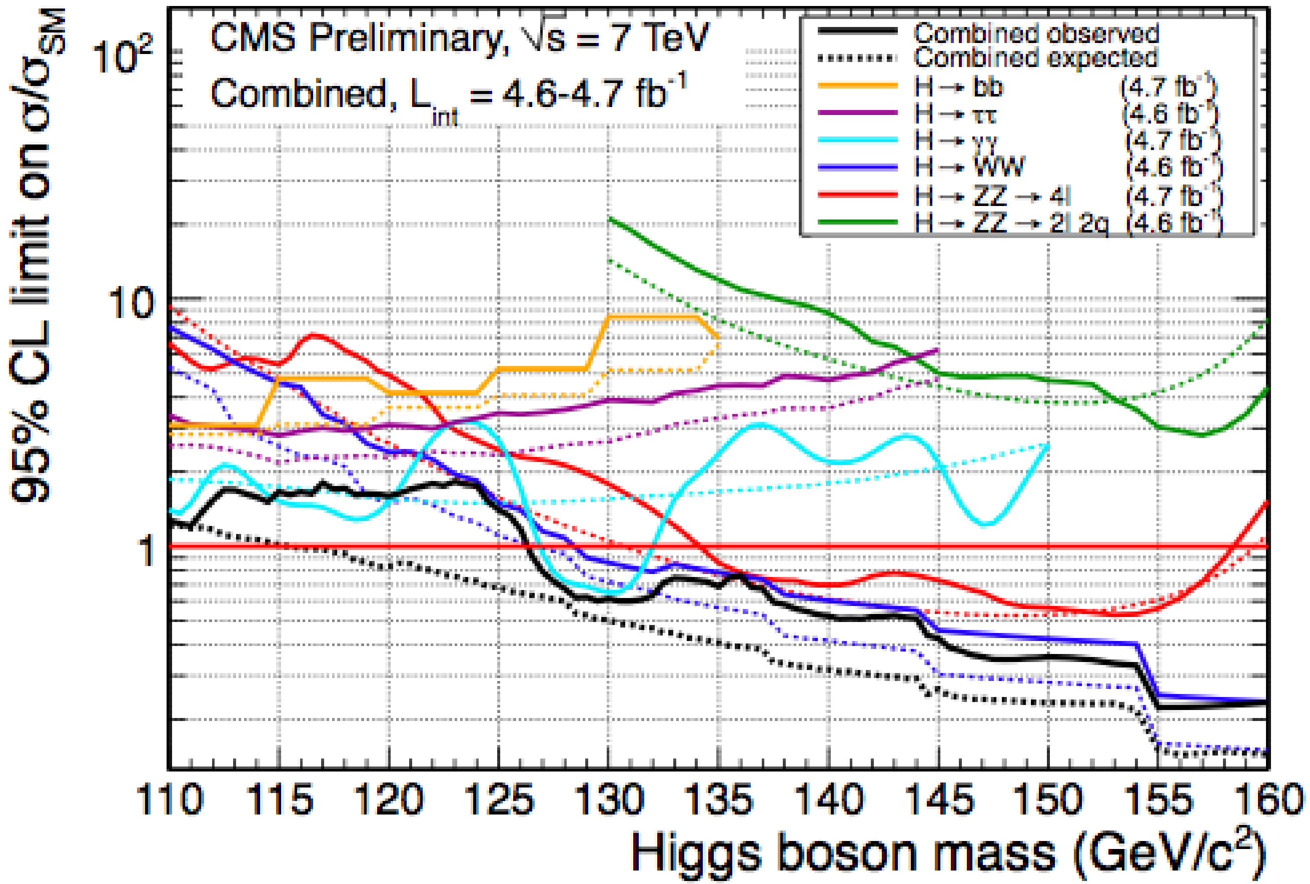


What is the true probability of a local excess ?

Trials factors estimations

- Number of independent searches being performed
  - Range of search in mass / Mass resolution
- Pseudo-data
  - Using toy MC to determine how often an excess as large can happen
- Approximation
  - For small P-values, in asymptotic regime, can count up-crossings of signal stength = 0, and determine global P-value from test statistic
    - In observed data

Approx: Trials factor = Range / resolution



Problem:
Concept of mass resolution not clear in an MVA
Also MVAs trained separately at each mass point !
(Mass points are correlated)

Not used by LHC

(At around 120 GeV)

H→γγ (1-3%)

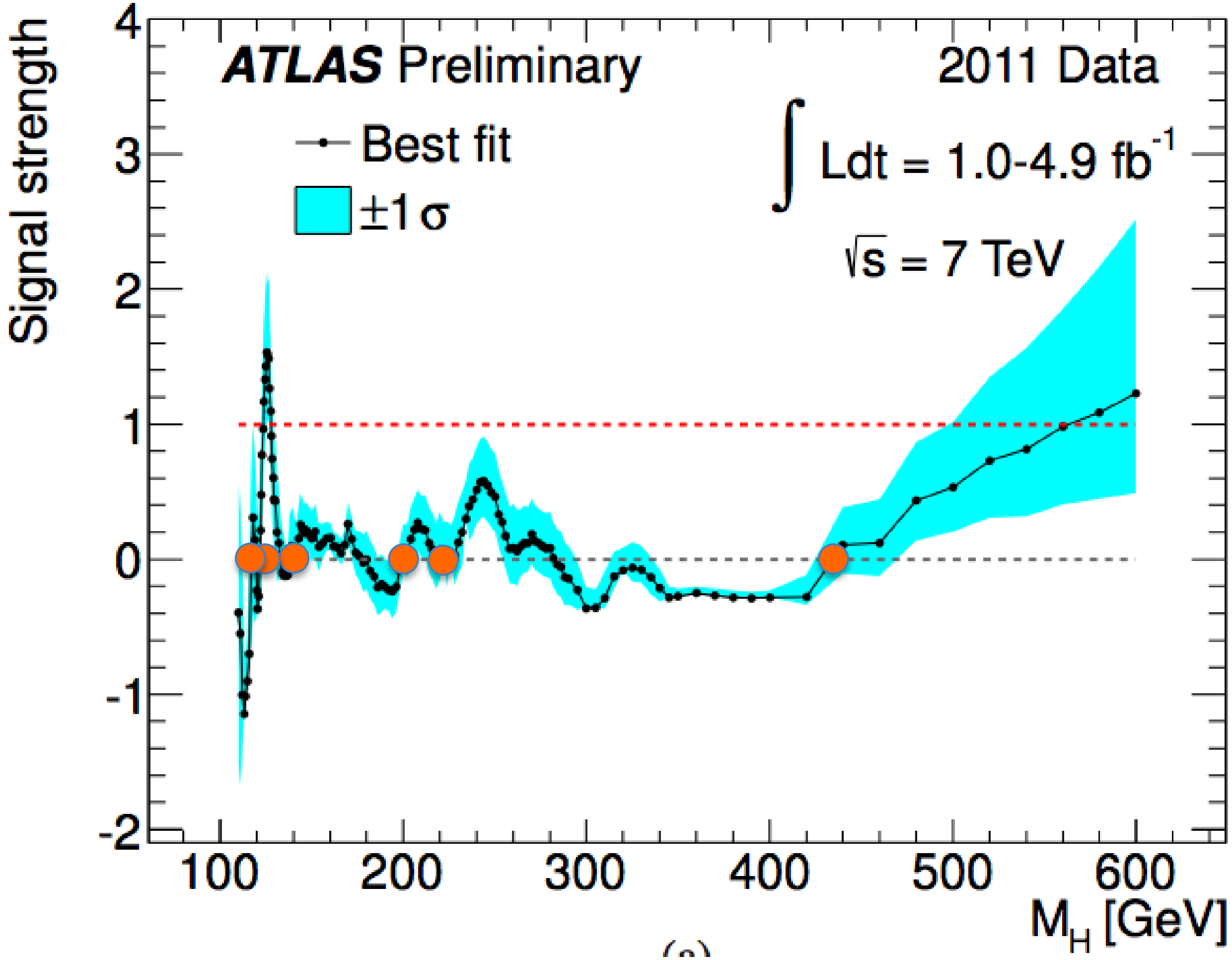
H→ZZ→4l (1-2%)

H→WW (20%)

H→ττ (20%)

H→bb (10%)

Trials factor: Up-crossings



At 0:
ATLAS: 6 up-crossings
CMS: 8 up-crossings

Trials factor not that sensitive to statistical Fluctuations

Cross-checked using different signal strengths and with statistical uncertainties

Significances of excesses

- ATLAS: 126 GeV
  - 3.6 Sigma local P-value
  - 2.2 Sigma with trials factor

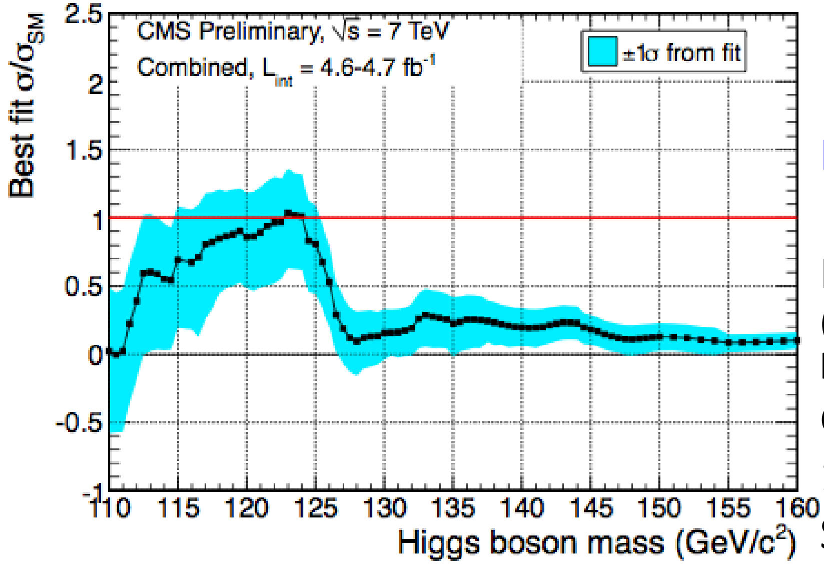
- CMS: 119 GeV
  - 2.6 Sigma local P-value
  - 0.6 Sigma with trials factor

But what is the right Look Elsewhere Effect ?

- CMS & ATLAS search 110 – 600 GeV
  - Decided a priori based on experimental reach
  - Generates a large Look Elsewhere Effect
  - Do we really expect a SM Higgs boson to be 600 GeV ?
- Could use previous experimental exclusions for prior
  - ATLAS uses 2fb^{-1} LHC combination to motivate restricted window
    - 110 – 146 GeV :
    - 3.6 σ local → 2.2 σ (full mass range) → 2.5 σ (restricted)
  - But *unfair* to use subset of data both to define search window and perform search

Restricted mass range

- CMS restricted mass range
  - Statistical uncertainty of up-crossing technique in observed data is limited
    - CMS finds 1 up-crossing in this mass range



If this were to be used,
Global significance = local significance:

Instead, use pseudo-experiments instead of up-crossings

Problem: H→WW uses mass-dependent MVAs
(correlations between mass points because backgrounds not the same at each search mass)
Generate pseudo-data from background model at m_H ~ 125 GeV
Safe assumption for 110 – 145 GeV due to M_{WW} resol.
So 110 – 145 GeV is used by CMS (~ same as ATLAS)

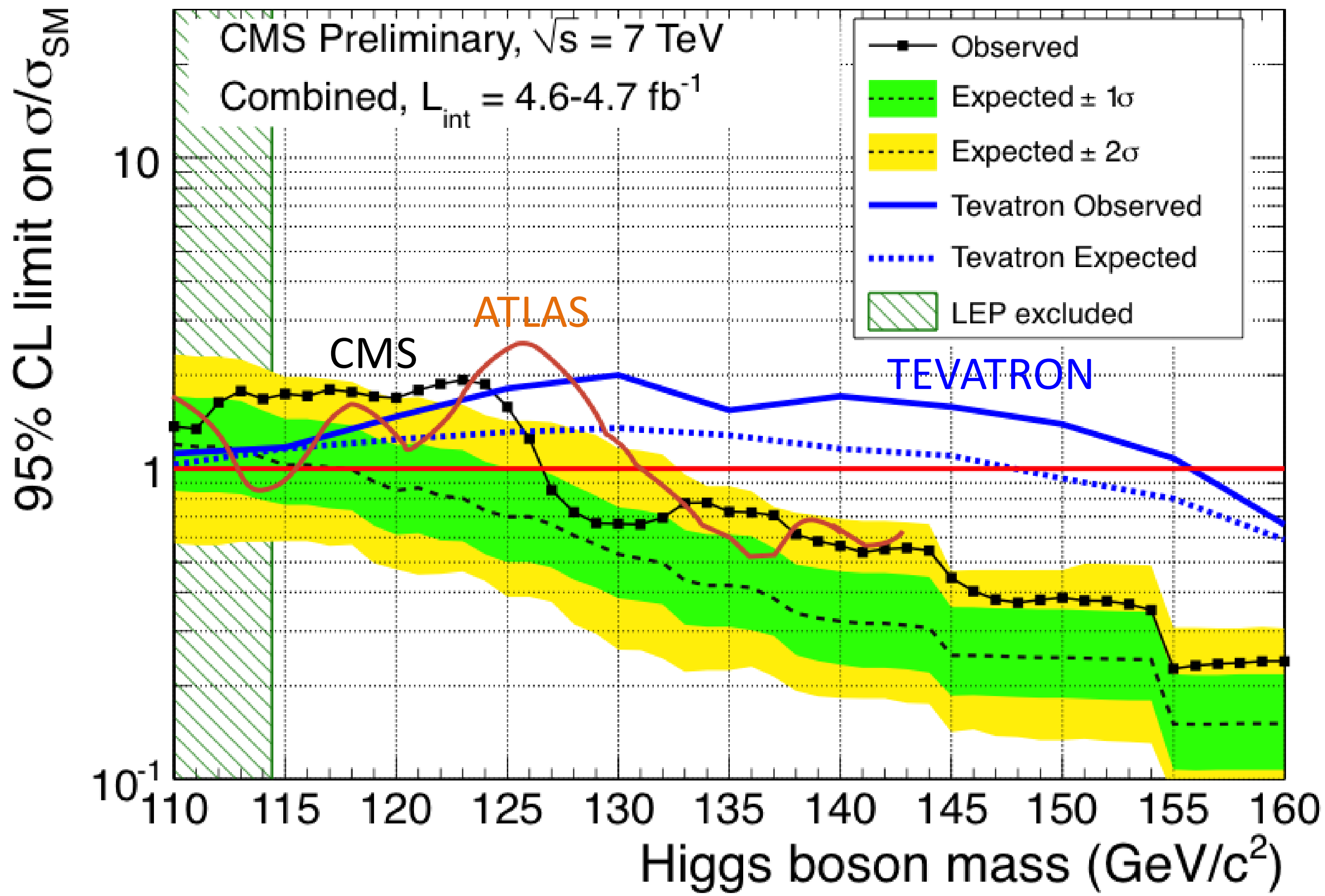
2.6 σ (local) → 0.6 σ (full mass range) → 1.9 σ (restricted)

What is the right restricted mass range ?

- CMS could use ATLAS' exclusion range for search window & vice versa
  - Not very agreeable since there are correlations between nuisance parameters of CMS & ATLAS
- Could split data into old data for exclusion, and new data for search window
  - Lots of work, and would be self-defeating since not all data would be used, reducing significance at the expense of reducing trials factor

- Instead … the SM Higgs boson is predicted by precision electroweak measurements (LEPEWWG)
  - $m_H < 161$ GeV at 95% CL
- So a more appropriate prior assumption for look elsewhere effect would be :
  - 110 – 161 GeV

Our picture of the Higgs boson



Conclusions from Higgs picture

- 

… but not forgotten
  - Higgs signal
    - Production at each mass
    - Production at each accelerator
    - Uncertainties
    - Uncertainties of exclusive final states
  - Higgs decays at each mass
  - Backgrounds at each mass
  - Backgrounds at each accelerator
  - Different statistical methods
  - Defining an excess

Higgs discovered

Rename

"God particle"

Newspaper Headline :
"Physicists prove God exists"

Rake in the $$$

Higgs excluded

Rename

"Devil particle"

Newspaper Headline :
"Physicists prove Devil does not exist"

$$$